

Perceptual learning, motor learning and automaticity

The neural basis of visual object learning

Hans P. Op de Beeck^{1*} and Chris I. Baker^{2*}¹ Laboratory of Biological Psychology, University of Leuven (K.U.Leuven), Tiensestraat 102, 3000 Leuven, Belgium² Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, 10 Center Drive, Building 10, 3N228, Bethesda, MD 20892, USA

Object vision in human and nonhuman primates is often cited as a primary example of adult plasticity in neural information processing. It has been hypothesized that visual experience leads to single neurons in the monkey brain with strong selectivity for complex objects, and to regions in the human brain with a preference for particular categories of highly familiar objects. This view suggests that adult visual experience causes dramatic local changes in the response properties of high-level visual cortex. Here, we review the current neurophysiological and neuroimaging evidence and find that the available data support a different conclusion: adult visual experience introduces moderate, relatively distributed effects that modulate a pre-existing, rich and flexible set of neural object representations.

The pervasive role of experience in visual processing

Sensory information processing in adult mammalian brains is highly malleable [1] with neural processing at all levels adapting to both the short- and long-term properties of the incoming information. In vision, prominent examples include short-term adaptation to input statistics in the retina [2], primary visual cortex (V1) [3] and subsequent cortical stages [4], and long-term reorganization due to changed visual input [5].

Nevertheless, cortical neural plasticity has often been viewed as more likely in visual regions selective for complex objects than in the input stage of processing, V1 [6]. It is unlikely that cortical representations could be constructed, *a priori*, to represent all possible objects that might be encountered throughout life. Indeed, human beings can recognize an almost infinite number of objects, most of us sharing the ability to individuate thousands of faces despite their similarity (two eyes, nose, and mouth in a standard configuration). Further, we often develop expertise in the recognition of particular types of objects such as cars, birds or wild mushrooms [7].

In apparent support of extensive learning processes, the past two decades have seen numerous electrophysiological and brain imaging studies reporting effects of learning on high-level visual representations. By learning we mean the effects of any form of experience, whether through passive

exposure or some sort of explicit training, such as learning to categorize [8] or discriminate [9] between objects. Most of these studies have focused on the lateral occipital, occipitotemporal and inferior temporal cortices in humans and the inferior temporal cortex in monkeys, which we collectively refer to as “IT cortex”. However, some studies have argued for a role of the whole cortical visual processing hierarchy in learning about objects and their features [10–12]. In this review, we critically evaluate the evidence and hypotheses about visual object learning in IT cortex in the context of the functional properties of these brain regions. We focus on long- rather than short-term learning effects (Box 1), in adults rather than during development (Box 2). We will conclude that, despite the common suggestion that object learning is supported by strong and often focal changes in neural representations [13,14], such as the possibility that a subpopulation of units tuned to complex images is created by and emerges due to a learning process [15,16], the empirical evidence suggests that these changes are moderate and distributed.

Visual object coding in IT cortex

Individual neurons and subregions in IT cortex are selective for complex or moderately complex stimuli, responding more strongly to some particular images or categories of images than to others [17–19]. The observed properties suggest that IT cortex contains explicit representations of stimulus dimensions such as shape, at least some visual categories and possibly even semantic categories, with a relative tolerance for changes in other stimulus characteristics such as position, size, lighting, and, to a lesser degree, clutter and viewpoint (see Figure 1). These representations are surprisingly versatile, with easy ‘read-out’ of object category and identity [20], while simultaneously containing information about other stimulus characteristics, for example, position [21,22].

Many studies of IT cortex have employed images of complex, real-life objects. With these images, it has proven difficult to determine the underlying dimensions that best describe the neural tuning functions (Figure 1). Other experiments have included parametric variations of shapes and have reported systematic tuning functions for a variety of shape dimensions [23–26], analogous to the tuning in V1 for more simple stimulus dimensions (e.g. grating orientation). However, such tuning is demonstrated in

Corresponding authors: Op de Beeck, H.P. (hans.opdebeeck@psy.kuleuven.be); Baker, C.I. (bakerchris@mail.nih.gov)

* These authors contributed equally to the work.

Box 1. Short-term and long-term learning

There are very good reasons to differentiate learning effects according to the time scale over which those effects are seen. The molecular mechanisms are very different for modifications over seconds and minutes compared to those over hours and days. For example, long-term changes have to involve protein synthesis and gene transcription. The distinction between short- and long-term has been made in the literature on perceptual learning of basic visual properties such as orientation selectivity, and there are a few prototypical examples. Typical short-term effects are visual after-effects in which the continuous presentation of one adapting stimulus (e.g., a leftward oriented line) changes the perception of another stimulus (e.g., a vertical line [81]) presented immediately after the adapting stimulus. Typical long-term effects are the decrease in orientation or texture discrimination thresholds that are induced gradually during several weeks of training [82,83]. At least some studies have suggested a role of sleep in these long-term paradigms [84,85], pointing to a qualitative difference between short-term and long-term learning in the underlying mechanisms. The literature of visual object learning also contains studies of short-term adaptation and multiple-day learning. Nevertheless, the distinction is less clear. Object adaptation and priming studies often include intervening stimuli, and such effects tend to integrate over many tens of stimulus presentations [86–88]. Is it appropriate to refer to this as a short-term effect? On the other hand, object learning studies in humans tend to include a relatively short training period of on average a few hours (ranging from less than one hour to at most ten one-hour sessions) (e.g. [9,89–93]), and the role of sleep in these paradigms has not been investigated. Is this a long-term effect? Whereas we can make an operational distinction between within-session and between-session learning, for now it is premature to attach too much weight to this distinction. Nevertheless, we will restrict our discussion to studies traditionally believed to focus on long-term learning effects, and refer to other reviews for a discussion of more short-term, within-session adaptation effects [94].

relatively restricted stimulus sets, and we are far from understanding the representational space of more complex images. In this review we will describe the stimulus selectivity of IT neurons and its modulation by learning in terms of tuning curves for dimensions and features, but this

Box 2. Developmental changes in object representations

Object representations in the adult brain are the result of a long developmental history that affects all levels of the visual processing hierarchy. Even in V1, many of the properties of neurons and maps, such as orientation selectivity, directional selectivity, and ocular dominance, depend on visual experience early in life (first few months after birth) [95,96]. After this time period, some complex response properties are already apparent in monkey IT cortex [97], but the physiological data in animals are not systematic enough to make quantitative comparisons of object representations as a function of age [98]. Evidence from several non-invasive studies suggests that face processing is the consequence of an interplay between face-specific innate mechanisms and visual experience [99]. Further, behavioral studies in human and monkey infants have revealed experience-dependent face-specific biasing mechanisms that are further tuned by visual experience [100,101]. Early visual experience has a special importance for face processing given that visual deprivation in the first few months of life disrupts the development of the holistic processing that is characteristic for face perception in normal adults [102]. Nevertheless, neural markers of category-specific object processing in the human brain change up to adolescence [103]. In sum, although plasticity during development is probably much higher than during adulthood, the currently available data suggest for both periods that visual experience introduces incremental effects that modulate a pre-existing set of neural object representations.

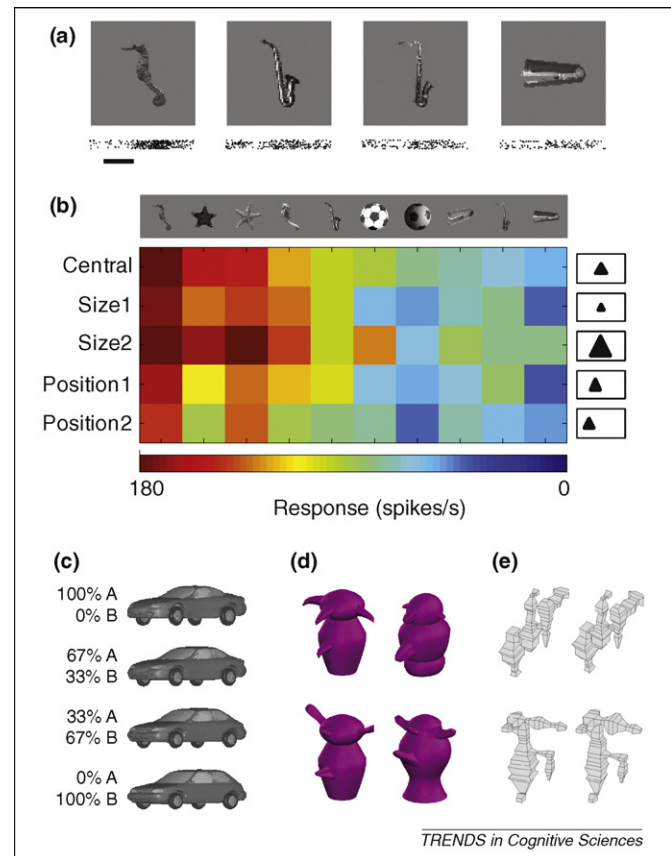


Figure 1. Selectivity for visual objects in IT cortex. (a) Responses of a single IT neuron to four complex images. The raster plots under each image show the occurrence of action potentials over time (each line corresponds to a single trial). The horizontal line below the panel represents the time interval over which the stimulus was presented (duration: 100 ms). With only the responses to these four images, it would be tempting to describe the variation in response strength in terms of overall shape orientation. However, the responses to a broader set of ten images (see panel b) reveal no obvious relationship between this or any other property of the images and the neural response. (b) Color plot showing the magnitude of response elicited from the same neuron by a larger set of ten images presented at various sizes and retinal positions. The responses of this neuron, and of many neurons in IT cortex, are not absolutely invariant under the different image transformations, but the rank order of the images as a function of response strength is remarkably similar across different transformations. This phenomenon has been referred to as relative invariance or tolerance [17]. Data in panels a,b were provided by Chou Hung from the dataset described in [20]. (c–e) Examples of images that have been used to study visual object learning. Other images are shown in Figures 3 and 4. (c) Morphed cars. A parametric stimulus space was created by morphing between two car exemplars (A and B) [58]. (d) Greebles. Like faces, all individual Greebles have the same number of parts in the same configuration [40]. (e) Complex polygonal objects. All objects were matched for low-level perceptual features but differed in the configuration of parts [74].

selectivity could also be characterized in terms of more abstract principles, such as the statistical structure of images [27,28].

Effects of learning on object representations:

Theoretical considerations

Learning and the single neuron

At the single-neuron level, we differentiate two tuning properties that might be modified by learning: the optimal stimulus, and the tuning function around that optimum. For example, consider a neuron that shows moderate tuning for two separate stimulus dimensions (Figure 2a–b). Learning might increase or decrease the selectivity across one or both dimensions, it might change the optimal

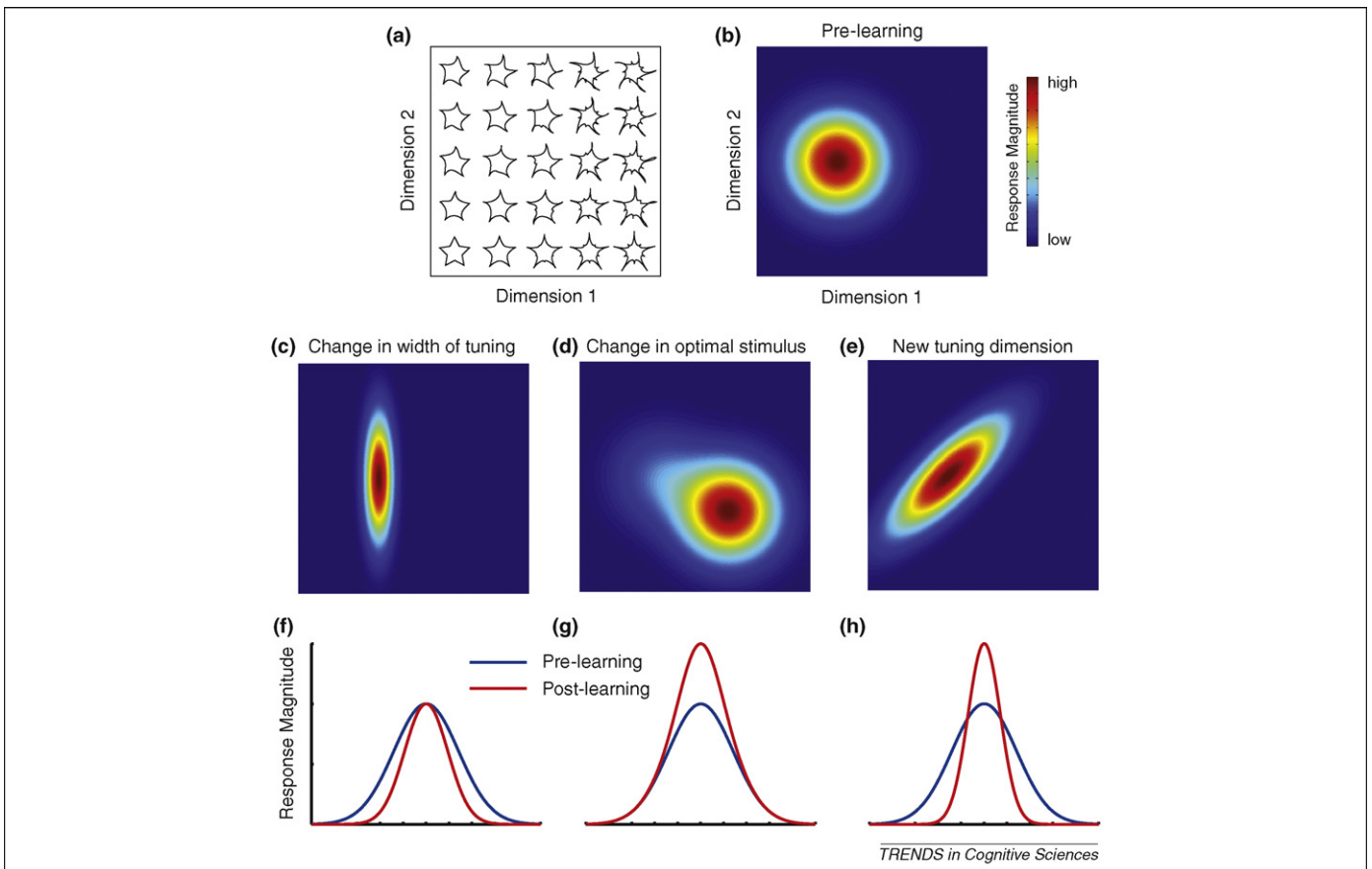


Figure 2. Possible effects of learning on tuning of individual neurons. (a) Illustrative stimulus space comprising two complex stimulus dimensions. (b) Prior to learning an individual neuron may show some degree of tuning for both dimensions with strong responses for some stimuli along each dimension and weaker responses for others. (c) Learning may alter the width of tuning along a given dimension. For example, learning to make fine-grained discriminations within Dimension 1 (relevant dimension) might lead to a sharpening of tuning (increased selectivity) along that dimension, and a broadening of tuning (decreased selectivity) for Dimension 2 (irrelevant dimension). (d) Learning may also change the optimal stimulus for a given neuron. For example, experience with stimuli from only the lower right part of the stimulus space might lead any neurons with pre-learning selectivity for other parts of the stimulus space (as shown here) to shift their optimal stimulus to the relevant part of the space. (e) Learning may also lead to the development of new tuning dimensions. For example, the relevant features for learning may not be encapsulated in any single dimension, but may require a conjunction across dimensions. In this case the optimal stimulus for any given value along one dimension may depend on the value of the other dimension, effectively creating a new dimension. (f-h) Illustration of how sharpening of tuning width might be achieved for a given trained dimension: (f) Reducing the response to less preferred stimuli while maintaining the response to the optimal stimulus; (g) Increasing the response to the most preferred stimuli while maintaining the (small) response to less preferred stimuli; or (h) Increasing the response to the optimal stimulus, but reducing the response to less preferred stimuli. Importantly, this latter mechanism could maintain the average response strength to all stimuli within the trained dimension.

stimulus, and/or it might change the dimensions that are encoded by the neuron (Figure 2c-e). Importantly, these changes in selectivity will interact with changes in response strength (Figure 2f-h). For example, a sharpening of tuning could be achieved either by reducing responses to non-preferred stimuli, by increasing responses to preferred stimuli, or a combination of both.

What factors are likely to be important in driving these sorts of changes? Whereas theories of coding in V1 propose an optimal representation of visual images given the local spatiotemporal statistical properties of natural images [29], theories of object coding typically propose that these representations reflect more global statistical properties: the dependencies between more distant image patches [30] and at a longer time scale [31]. Many of these spatial dependencies might be picked up by bottom-up Hebbian learning mechanisms sensitive for conjunctions of features [32], possibly augmented by mechanisms sensitive to probability statistics [33]. The temporal dependencies might require a similar learning mechanism with a memory trace [34]. In addition to these bottom-up mechanisms, the

properties of object representations must also be sensitive to top-down information about which distinctions between images are most relevant or informative. Recent computational modeling emphasizes that tuning for moderately complex features would be the optimal compromise between the bottom-up statistics and the top-down requirements imposed on object representations [35].

Learning and the population

At the population level there are two important aspects to consider: sparseness and clustering [36]. Sparseness refers to the distribution of learning changes across the population of IT neurons. At one extreme, learning could modify the response properties of only a small number of neurons. At the other extreme, learning could modify the properties of all neurons. Clustering, on the other hand, refers to the spatial organization of learning-related changes. Learning could increase clustering of neurons with similar selectivity, perhaps optimizing interactions between neurons.

If learning effects are sparse, then which factors determine the neurons that will be modified most by learning?

At least three factors have been proposed. First, learning might specifically involve a small subset of neurons that prior to learning showed limited responsiveness [13]. After training these initially unresponsive neurons would be tuned to complex novel objects such as ‘paperclips’ [13,37,38], requiring both sharpening of tuning (Figure 2c) and a shift in the optimal stimulus (Figure 2d). Second, learning might specifically involve face-selective neurons and regions [39,40]. This prediction is rooted in the hypothesis that expertise is the major cause for the selectivity of face-selective regions. Before training, these brain regions would be characterized by strong face selectivity, but not necessarily by a response to the to-be-trained objects. Third, studies of orientation discrimination learning in retinotopic areas have suggested that learning might modify the tuning of the neurons that are most informative for solving the discrimination task [41], and a similar mechanism could be at work during object learning.

To understand visual object learning, it is critical to consider the degree of sparseness and clustering and the factors that determine which neurons are most affected. With the limited sampling afforded by single unit recording, it would be very easy to miss sparse learning-related changes, even if the few modified neurons show very large changes in tuning properties. Similarly, given the spatial resolution of functional imaging methods, it might only be possible to detect distributed learning-related changes and even then, only when there is significant clustering of those changes.

Linking theory to empirical data

We will show below that the current empirical evidence is consistent with general predictions from the computational proposals: learning changes neural tuning and responsiveness according to both bottom-up stimulus characteristics and top-down task constraints. However, almost no studies have been computationally motivated to investigate more specific hypotheses, so the link between the empirical findings and the theories is tenuous.

There is one important caveat to the learning mechanisms discussed above. The diverse selectivity observed in IT cortex suggests that for any given task some neurons will be more informative than others. One possible mechanism underlying object learning might be the optimization of the read-out of the most informative IT neurons [42–44], as has also been proposed in the domains of orientation and motion learning [45,46], without any changes of the response properties within IT cortex.

Effects of learning on object representations: Empirical evidence

Experiments typically compare neural responses to learned objects with an unlearned baseline, either obtained before learning [9,47], after learning to a set of unlearned objects [48,49] or in different, ‘naive’ subjects [50,51]. Importantly, studies do not typically track changes in the properties of single neurons across days and weeks because of technical limitations, and longitudinal studies have been limited to 1–2 hours at most [52–54]. fMRI allows changes to be tracked across long periods of time

in individual voxels, but each voxel contains hundreds of thousands of neurons. So, all inferences are necessarily based on the comparison of population statistics between experimental and baseline conditions.

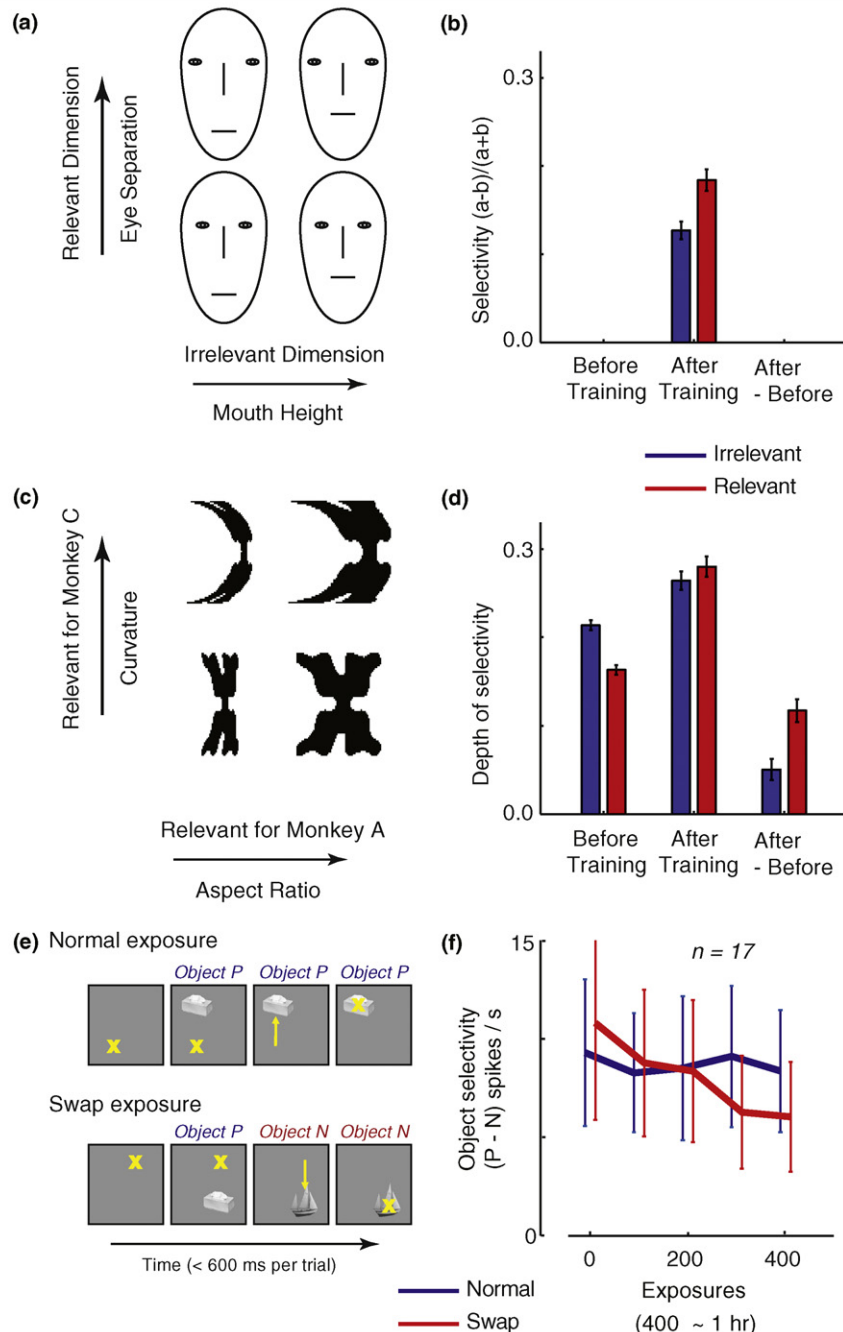
Learning and the single neuron

Learning changes the selectivity or tuning for experienced objects, with or without an additional role for task constraints. In monkeys, some studies have reported a strong increase in selectivity for trained compared to untrained objects [38,50] as well as an increased selectivity for relevant compared to irrelevant stimulus dimensions [8] (Figure 3a,b). Later studies, which controlled for pre-existing selectivity biases and spatial attention, reported much smaller effects, often only a few spikes per second [47,48,55,56] (Figure 3c,d, 4a,b). Using an indirect fMRI adaptation method to determine average changes in selectivity in a neural population [57–59], human studies have confirmed a general increase in selectivity for trained objects in object-selective cortex, but without a difference in selectivity between relevant and irrelevant dimensions. Overall, effects have been found consistently by comparing trained with untrained objects, but more specific effects such as differences between relevant and irrelevant dimensions seem to be small and harder to detect.

Studies have also reported decreased selectivity for stimuli that become associated or predictive of each other [60,61]. The mechanism behind such associative coding might underlie the creation of tolerance for image transformations such as position or orientation changes [34]. However, the effects have mostly been reported in multimodal cortical regions (perirhinal and entorhinal cortex) rather than unimodal visual areas (area TE in monkeys), appear weaker in TE [62] and are dependent on the integrity of perirhinal and entorhinal cortex [63]. Thus, there is ambiguity as to whether associative coding is an instance of visual object learning or of semantic memory.

While the temporal association in the classic paired-associate task is artificial, other studies have used temporal dynamics during training that resemble natural vision. For example, a comparison of view invariance between familiar and unfamiliar objects, where the familiar objects had been placed in the monkey’s living environment, suggested stronger invariance for the familiar objects [64]. However, too few neurons were tested in the critical comparisons to allow strong conclusions. A recent within-session longitudinal study induced associations between objects with temporal dynamics that come close to those encountered in free viewing, and demonstrated that tolerance can be modulated in this situation [52] (Figure 3e,f). Nevertheless, it remains unclear whether these findings reflect the same mechanisms that are responsible for the degree of tolerance typically observed in IT cortex.

As noted earlier, learning might not just modify the degree of selectivity, but also the optimal stimuli and tuning dimensions (Figure 2d,e). Some studies have provided circumstantial evidence for such changes. For example, Logothetis and colleagues found neurons in monkey IT cortex with a strong preference for trained views of complex paperclip objects, but noted that “no



TRENDS in Cognitive Sciences

Figure 3. Studies of the effect of a specific training or exposure regime on the responses of neurons in monkey IT cortex. These studies illustrate the role of top-down feedback about task relevance (**a–d**) as well as the role of bottom-up spatiotemporal statistics of the visual input (**e–f**). **(a)** Subset of the face stimuli used in [8], with an indication of the relevance of the stimulus dimensions for the behavioral task that monkeys performed during training and recordings. The relevant stimulus dimensions were the same for all monkeys. **(b)** Selectivity for differences along relevant and irrelevant face dimensions (subset is shown in panel **(a)**) as recorded after training, showing greater selectivity for the relevant than irrelevant dimension. Data were extracted from Figure 4 in [8]. **(c)** Subset of the shape contours used in [47]. One of two dimensions was relevant for the behavioral task that monkeys performed during training and during the recordings after training, and task relevance was counterbalanced across monkeys. **(d)** Selectivity for differences along relevant and irrelevant shape dimensions (illustrated in panel **(c)**) as recorded before training (during passive fixation) and after training. Note that the index of selectivity here is different from that shown in **(b)**. The third pair of bars displays the selectivity after training minus the selectivity before training. This study illustrates how important it is to know about pre-training selectivity in order to interpret after-training selectivity. Only after correcting for the pre-training selectivity at the population level do the results in panel **(d)** confirm the tentative conclusion from panel **(b)** that behavioral relevance during training matters for the degree of selectivity in IT cortex, at least when recordings are performed during the execution of the trained task. Several other studies, some of which used a different task context during recordings, have suggested little difference between relevant and irrelevant stimulus differences [55,57,58], but an overall increase in selectivity. **(e)** In another study [52], monkeys were exposed to two conditions during free viewing, one condition in which the screen display contained the same objects before and after a saccade (normal exposure), and another condition in which an object was changed during the saccade (swap exposure). The yellow 'x' refers to the position of fixations, and the yellow arrow to a saccade. Primates do not notice the transient on the screen caused by the swap because visual input is suppressed during a saccade. **(f)** The results from recordings performed after every 100 exposures in the design shown in panel **(e)**. Here we plot the difference in response to the initially preferred object image (object P) and the initially nonpreferred object image (object N) of each neuron. This response difference is not altered in the normal exposure condition, but it is reduced over time in the swapped exposure condition. So the two objects that are associated in time become more similar in terms of the neural responses they evoke.

selective responses were ever encountered for those views that the animal systematically failed to recognize” [38] (p. 558). Similarly, other studies [50,65] have also highlighted subpopulations of neurons with particularly high and selective responses for trained objects, but not for novel objects. However, none of these studies provided sufficient control conditions to firmly establish the nature of the underlying effects. Perhaps the best evidence for at least some minor changes in tuning dimensions comes from a study in which monkeys were trained to categorize multi-part baton stimuli based on the combination of parts present [48]. Comparing the patterns of selectivity for trained and untrained batons revealed greater coding for the specific combination of parts, rather than the presence of individual parts, in trained batons. With single-part properties as the original dimensions, the coding for part combinations is an example of the generation of a new tuning dimension reflecting the interaction between parts (as illustrated in Figure 2e).

Learning and the population

The simplest characteristic at the population level is the overall response in experimental and baseline conditions.

Some studies in monkeys have reported an increased population response [38,50,65–67], other studies a decreased response [48,49,68], other studies little or no effect [69]. Human imaging experiments suggest that both increases and decreases in response magnitude may occur across distributed areas of cortex [9,10,70,71]. Given the theoretical considerations above, it is not surprising that both increases and decreases have been observed - the results may depend on the initial tuning of the neurons with respect to the learned stimuli. The disparity across physiology studies probably also reflects the limited and non-uniform sampling (a few hundred neurons at best). Further, if the effects of learning are sparse with limited clustering, it will be difficult to isolate them.

These considerations bring us to the question of how sparse training effects are. Plotting the distribution of effects of interest, such as selectivity for trained and untrained objects, often reveals a general shift in this distribution (Figure 4b,e). However, this does not necessarily imply all neurons are affected equally by learning. Without tracking single neurons over time it is very difficult to establish the nature and distribution of learning effects. Nevertheless, fMRI has already revealed that

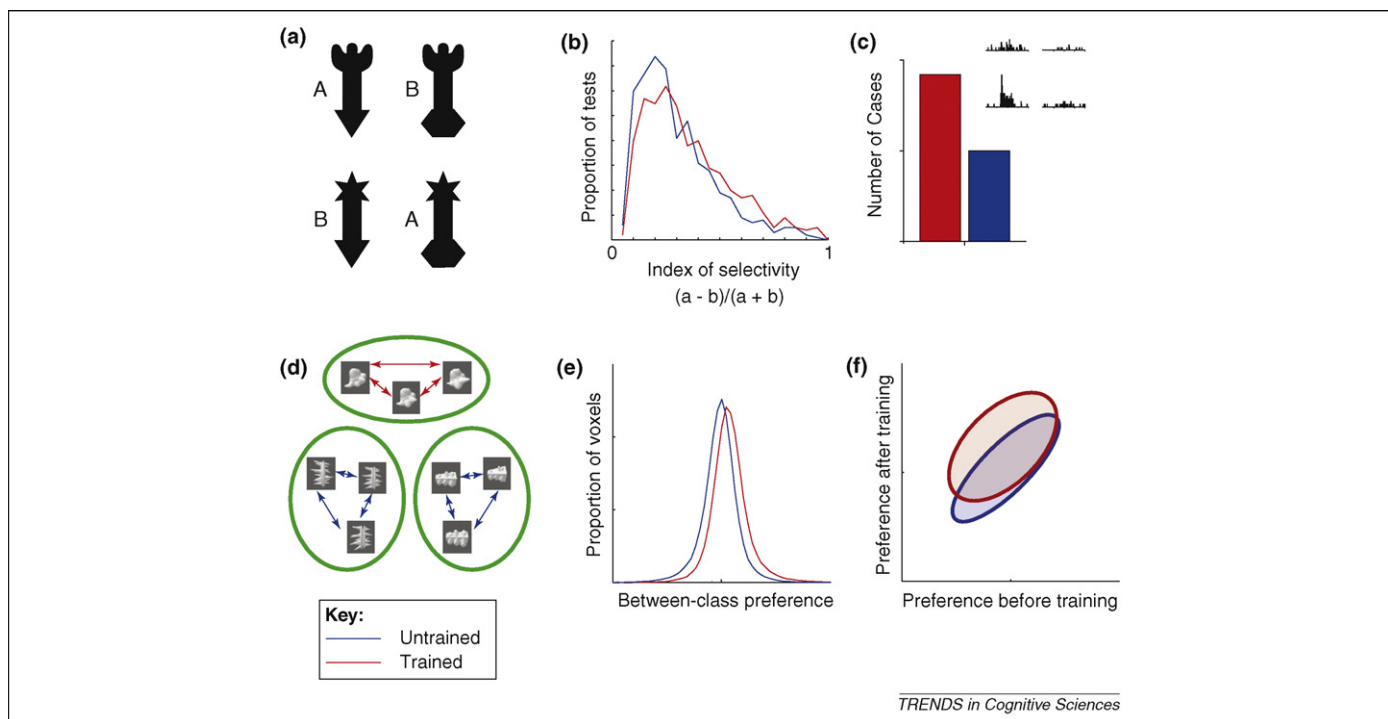


Figure 4. Studies of the distribution of learning effects across a neural population with and without knowledge of the neuronal properties before learning. (a) Examples of the multipart baton stimuli from ref. [48]. Monkeys were trained to respond to the batons according to the combination of parts present. (b) Distribution of selectivity for both trained (red) and untrained (blue) batons recorded after training. Note that the distribution for trained batons appears shifted to the right suggesting a full distribution of training effects rather than the emergence of a few highly selective neurons. (c) Bar plot showing the number of cases of coding for combinations of parts (as assessed by an interaction effect in an ANOVA). Almost twice as many instances of whole-object coding were observed for trained (red) compared with untrained (blue) batons. The inset shows the firing of an example neuron, responding to the four stimuli shown in (a). (d) In the fMRI study described in [9], each subject was trained to discriminate exemplars in one out of three object classes (here this class is shown on top, red arrows denote trained stimulus differences for one subject). Responses to the three object classes were measured in two scan sessions, one before and one after the training procedure. (e) Measured distribution of the response to the trained object class (red) relative to the response to the untrained object classes (blue). Responses are expressed as preferences by subtracting the mean response to untrained objects. These data are obtained after training. These histograms are calculated based on all visually active voxels in nine subjects, with a total of more than 100,000 voxels. Note that this display suggests a full distribution of the training effects, similar to the distributions shown in panel b. (f) Here we make use of the knowledge about pre-training responses that was obtained in the fMRI study by scanning prior to any training. The preference to the trained and untrained object classes after training (data from second scan session) is plotted as a function of the preference measured before training. For clarity we do not show all individual voxels (see Figure 6 in [9]), but a simplification of the two distributions with two elliptic surfaces (trained, red; untrained, blue). The response was higher for the trained objects post-training (as shown by the vertical shift in the red ellipse relative to the blue ellipse; the same effect is visible in the marginal distributions which were shown in panel e). Further, the between-session correlation was lower for trained than untrained objects, (as shown by the fatter ellipse - the ratio of long versus short axis equals $\sqrt{1+r}/\sqrt{1-r}$, with r the between-session correlation after correction for the within-session reliability of the data). Thus, training changed the pattern of selectivity across cortex, which indicates the effect of training varied between voxels and was not fully distributed.

training changes the pattern of selectivity across voxels [9] (Figure 4f), indicating heterogeneous training effects across voxels that are not fully distributed. One other fMRI study has suggested relatively focal effects of learning to read a particular alphabet (Hebrew) in one part of IT cortex, the visual word-form area [51]. In addition, a single-unit study in monkeys has suggested that visual experience increases the clustering of neurons in perirhinal cortex with similar response properties [72], which should also change and even enhance the pattern of selectivity as measured with fMRI.

Since the training effects are not fully distributed, then the question is: which neurons will be affected most? In the theoretical section we mentioned three possible factors. Almost all empirical efforts so far have focused on the role of face selectivity and some fMRI studies have indeed found strong training effects in face-selective cortex [39,40,73]. However, most of these studies did not systematically compare face-selective cortex with other regions in IT cortex. Other, more recent studies have shown that relative moderate learning effects are distributed throughout IT cortex [9,58,74,75], without any relationship to face selectivity.

Another candidate factor is how informative a neuron is for the task at hand. This hypothesis is supported by studies of the effect of orientation discrimination learning in retinotopic areas V1-V4 [41]. The preferred orientation of a neuron turned out to be a strong predictor of the strength of training effects. Without reference to the preferred orientation of neurons, these studies of orientation discrimination would support similar conclusions to the IT data: small effects distributed across a broad neural population (e.g., compare Figure 4b with Figure 3 in ref. [41]). The problem with applying this approach to the learning of more complex objects is that, as described above, visual objects occupy a multi-dimensional space of which the dimensions are not very well known. One potential way out of this problem would be to abandon the notion of an explicit representation of features or dimensions, and describe the tuning curves of neurons and experience-related changes in terms of relatively abstract notions of the statistical properties of complex visual images [27,28].

Nevertheless, ‘informativeness’ is a promising candidate to encompass all existing findings about the neural basis of visual object learning, and is closely related to a formal computational model [35,76]. Furthermore, it can also explain at which level in the visual processing hierarchy the effects would be most abundant: the level at which the selectivity of neurons fits best with the task at hand [50,77]. Finally, it might also explain why, sometimes, neural training effects are small despite large behavioral effects. Indeed, given the diverse tuning properties observed in IT cortex, learning may merely modify the read-out of this rich, ‘informative’ neural population [42,43].

Concluding remarks

Despite the prevalent view that IT cortex is highly plastic, the current evidence remains limited. Well-controlled studies find relatively small effects that seem to be widely distributed. The evidence is strong enough to uphold the

Box 3. Outstanding Questions

- What are the relative roles of bottom-up image statistics and feedback about task relevance for the size and the nature of learning-dependent changes in neural object representations? We mentioned that some, but not all, of the studies that focused on task-related effects described significant effects. However, most of the learning effects are seen in studies in which task relevance and mere exposure go hand in hand: the trained objects are task relevant as well as frequently encountered [9,38,48–50]. Disentangling these two factors is an important task for future studies.
- What is the relationship between the pre-existing object representations and the distribution and size of learning effects? Many questions relate to this general question: Are the most informative neurons changed the most? How much of the strong behavioral effects are due to changes in object representations in contrast to changes in how object representations are ‘read out’? Some of these questions can be partially answered with comparisons at the population level, but the introduction of computationally motivated studies as well as longitudinal studies that track the responses of single neurons for days and weeks will undoubtedly be an important step.
- How do object representations change during development? Are these changes qualitatively similar to what happens during visual learning in adults? Are there critical periods and how are these different for different visual properties?
- All primates with normal development and no brain damage seem to be capable of recognizing objects with some tolerance for various image transformations. What are the minimal requirements of the visual input and the minimal structural properties of the system needed to sustain the development of this behavioral capacity and of the underlying object representations in IT cortex?
- What are the synaptic processes underlying plasticity in IT cortex (e.g., long-term potentiation and depression)? Can these processes provide a rationale to separate short-term, within-session adaptation effects from longer-term, between-session learning effects?

view that learning modulates at least some aspects of object encoding in IT cortex, but more detailed questions remain unanswered (Box 3). Future studies need to combine computational models with empirical experiments in order to predict when the available representations are sufficient for task performance (no further changes necessary), and when not. Furthermore, studies should relate effects of learning to the specific properties of neurons in order to pinpoint the potentially small sub-population of IT neurons that are targeted by learning. This sub-population might be defined by tuning properties (e.g., preferred objects or features), type of neurons (e.g., inter-neurons), or anatomical position relative to the organizational units in IT cortex (e.g., columns and patches). Answering such questions would be much easier if learning studies could track the response properties of single neurons over days and weeks. This methodology is currently being developed [78–80].

We have highlighted one specific hypothesis, that the strength of learning effects might be related to the pre-learning usefulness of neurons for the learned task and stimuli. This hypothesis seems quite obvious, given that learning effects are intimately bounded by the nature of the pre-learning representation. At the same time, it unites many benefits: it provides a link to computational arguments [35], and it offers a clear view on the expected strength of learning effects given the already existing representations. However, to test this

hypothesis convincingly, further technical developments are necessary to track single-neuron responses over long periods of time.

Acknowledgments

We thank Annie Chan, Assaf Harel, Dwight Kravitz, Sue-Hyun Lee and Rufin Vogels for comments on the manuscript, and Wouter De Baene, James DiCarlo, Chou Hung, Xiong Jiang, Nuo Li, Christopher Moore, Charan Ranganath, Maximilian Riesenhuber and Rufin Vogels for providing stimuli and data values. Support was provided by the Human Frontier Science Program (grant CDA 0040/2008 to H.O.d.B.), the Fund for Scientific Research - Flanders (grant 1.5.022.08 to H.O.d.B.), and the NIH Intramural Research Program at NIMH.

References

- Buonomano, D.V. and Merzenich, M.M. (1998) Cortical plasticity: from synapses to maps. *Annu. Rev. Neurosci.* 21, 149–186
- Hosoya, T. *et al.* (2005) Dynamic predictive coding by the retina. *Nature* 436, 71–77
- Gutnisky, D.A. and Dragoi, V. (2008) Adaptive coding of visual information in neural populations. *Nature* 452, 220–224
- Kohn, A. and Movshon, J.A. (2004) Adaptation changes the direction tuning of macaque MT neurons. *Nat. Neurosci.* 7, 764–772
- Dilks, D.D. *et al.* (2007) Human adult cortical reorganization and consequent visual distortion. *J. Neurosci.* 27, 9585–9594
- Gilbert, C.D. *et al.* (2001) The neural basis of perceptual learning. *Neuron* 31, 681–697
- Bukach, C.M. *et al.* (2006) Beyond faces and modularity: the power of an expertise framework. *Trends Cogn. Sci.* 10, 159–166
- Sigala, N. and Logothetis, N.K. (2002) Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* 415, 318–320
- Op de Beeck, H.P. *et al.* (2006) Discrimination training alters object representations in human extrastriate cortex. *J. Neurosci.* 26, 13025–13036
- Kourtzi, Z. *et al.* (2005) Distributed neural plasticity for shape learning in the human visual cortex. *PLOS Biol.* 3, e204
- Sigman, M. *et al.* (2005) Top-down reorganization of activity in the visual pathway after learning a shape identification task. *Neuron* 46, 823–835
- Hochstein, S. and Ahissar, M. (2002) View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron* 36, 791–804
- Sheinberg, D. and Logothetis, N.K. (2002) Perceptual learning and the development of complex visual representations in temporal cortical neurons. In *Perceptual Learning* (Fahle, M. and Poggio, T., eds), pp. 95–124, MIT Press
- Gauthier, I. and Logothetis, N. (2000) Is face recognition not so unique, after all? *Cognit. Neuropsychol.* 17, 125–142
- Riesenhuber, M. and Poggio, T. (1999) Hierarchical models of object recognition in cortex. *Nat. Neurosci.* 2, 1019–1025
- Huxlin, K.R. (2008) Perceptual plasticity in damaged adult visual systems. *Vision Res.* 48, 2154–2166
- DiCarlo, J.J. and Cox, D.D. (2007) Untangling invariant object recognition. *Trends Cogn. Sci.* 11, 333–341
- Tsao, D.Y. and Livingstone, M.S. (2008) Mechanisms of face perception. *Annu. Rev. Neurosci.* 31, 411–437
- Tanaka, K. (2003) Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but slightly different stimulus selectivities. *Cereb. Cortex* 13, 90–99
- Hung, C.P. *et al.* (2005) Fast readout of object identity from macaque inferior temporal cortex. *Science* 310, 863–866
- Op de Beeck, H. and Vogels, R. (2000) Spatial sensitivity of macaque inferior temporal neurons. *J. Comp. Neurol.* 426, 505–518
- Schwarzlose, R.F. *et al.* (2008) The distribution of category and location information across object-selective regions in human visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 105, 4447–4452
- Kayaert, G. *et al.* (2005) Tuning for shape dimensions in macaque inferior temporal cortex. *Eur. J. Neurosci.* 22, 212–224
- Brincat, S.L. and Connor, C.E. (2004) Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nat. Neurosci.* 7, 880–886
- Op de Beeck, H. *et al.* (2001) Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. *Nat. Neurosci.* 4, 1244–1252
- Yamane, Y. *et al.* (2008) A neural code for three-dimensional object shape in macaque inferotemporal cortex. *Nat. Neurosci.* 11, 1352–1360
- Schwarzkopf, D.S. *et al.* (2009) Flexible learning of natural statistics in the human brain. *J. Neurophysiol.* 102, 1854–1867
- Schwarzkopf, D.S. and Kourtzi, Z. (2008) Experience shapes the utility of natural statistics for perceptual contour integration. *Curr. Biol.* 18, 1162–1167
- Simoncelli, E.P. and Olshausen, B.A. (2001) Natural image statistics and neural representation. *Annu. Rev. Neurosci.* 24, 1193–1216
- Fiser, J. and Aslin, R.N. (2005) Encoding multielement scenes: statistical learning of visual feature hierarchies. *J. Exp. Psychol.* 134, 521–537
- Wallis, G. and Bulthoff, H.H. (2001) Effects of temporal association on recognition memory. *Proc. Natl. Acad. Sci. U. S. A.* 98, 4800–4804
- Mel, B.W. and Fiser, J. (2000) Minimizing binding errors using learned conjunctive features. *Neural. Comput.* 12, 731–762
- Orban, G. *et al.* (2008) Bayesian learning of visual chunks by human observers. *Proc. Natl. Acad. Sci. U. S. A.* 105, 2745–2750
- Wallis, G. and Rolls, E.T. (1997) Invariant face and object recognition in the visual system. *Prog. Neurobiol.* 51, 167–194
- Ullman, S. (2007) Object recognition and segmentation by a fragment-based hierarchy. *Trends Cogn. Sci.* 11, 58–64
- Reddy, L. and Kanwisher, N. (2006) Coding of visual objects in the ventral stream. *Curr. Opin. Neurobiol.* 16, 408–414
- Logothetis, N.K. and Pauls, J. (1995) Psychophysical and physiological evidence for viewer-centered object representations in the primate. *Cereb. Cortex* 5, 270–288
- Logothetis, N.K. *et al.* (1995) Shape representation in the inferior temporal cortex of monkeys. *Curr. Biol.* 5, 552–563
- Gauthier, I. *et al.* (2000) Expertise for cars and birds recruits brain areas involved in face recognition. *Nat. Neurosci.* 3, 191–197
- Gauthier, I. *et al.* (1999) Activation of the middle fusiform ‘face area’ increases with expertise in recognizing novel objects. *Nat. Neurosci.* 2, 568–573
- Raiguel, S. *et al.* (2006) Learning to see the difference specifically alters the most informative V4 neurons. *J. Neurosci.* 26, 6589–6602
- Ghose, G.M. *et al.* (2002) Physiological correlates of perceptual learning in monkey V1 and V2. *J. Neurophysiol.* 87, 1867–1888
- Op de Beeck, H.P. *et al.* (2008) The representation of perceived shape similarity and its role for category learning in monkeys: a modeling study. *Vision Res.* 48, 598–610
- Li, S. *et al.* (2007) Flexible coding for categorical decisions in the human brain. *J. Neurosci.* 27, 12321–12330
- Petrov, A.A. *et al.* (2005) The dynamics of perceptual learning: an incremental reweighting model. *Psychol. Rev.* 112, 715–743
- Law, C.T. and Gold, J.I. (2008) Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat. Neurosci.* 11, 505–513
- De Baene, W. *et al.* (2008) Effects of category learning on the stimulus selectivity of macaque inferior temporal neurons. *Learn. Mem.* 15, 717–727
- Baker, C.I. *et al.* (2002) Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nat. Neurosci.* 5, 1210–1216
- Freedman, D.J. *et al.* (2006) Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cereb. Cortex* 16, 1631–1644
- Kobatake, E. *et al.* (1998) Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *J. Neurophysiol.* 80, 324–330
- Baker, C.I. *et al.* (2007) Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proc. Natl. Acad. Sci. U. S. A.* 104, 9087–9092
- Li, N. and DiCarlo, J.J. (2008) Unsupervised natural experience rapidly alters invariant object representation in visual cortex. *Science* 321, 1502–1507
- Messinger, A. *et al.* (2001) Neuronal representations of stimulus associations develop in the temporal lobe during learning. *Proc. Natl. Acad. Sci. U. S. A.* 98, 12239–12244

- 54 Jagadeesh, B. *et al.* (2001) Learning increases stimulus salience in anterior inferior temporal cortex of the macaque. *J. Neurophysiol.* 86, 290–303
- 55 Freedman, D.J. *et al.* (2003) A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *J. Neurosci.* 23, 5235–5246
- 56 Cox, D.D. and DiCarlo, J.J. (2008) Does learned shape selectivity in inferior temporal cortex automatically generalize across retinal position? *J. Neurosci.* 28, 10045–10055
- 57 Gillebert, C.R. *et al.* (2009) Subordinate categorization enhances the neural selectivity in human object-selective cortex for fine shape differences. *J. Cogn. Neurosci.* 21, 1054–1064
- 58 Jiang, X. *et al.* (2007) Categorization training results in shape- and category-selective human neural plasticity. *Neuron* 53, 891–903
- 59 van der Linden, M. *et al.*, (2009) Formation of Category Representations in Superior Temporal Sulcus. *J. Cogn. Neurosci.* (Early Access; doi:10.1162/jocn.2009.21270)
- 60 Miyashita, Y. (1988) Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature* 335, 817–820
- 61 Sakai, K. and Miyashita, Y. (1991) Neural organization for the long-term memory of paired associates. *Nature* 354, 152–155
- 62 Naya, Y. *et al.* (2003) Forward processing of long-term associative memory in monkey inferotemporal cortex. *J. Neurosci.* 23, 2861–2871
- 63 Higuchi, S. and Miyashita, Y. (1996) Formation of mnemonic neuronal responses to visual paired associates in inferotemporal cortex is impaired by perirhinal and entorhinal lesions. *Proc. Natl. Acad. Sci. U. S. A.* 93, 739–743
- 64 Booth, M.C. and Rolls, E.T. (1998) View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cereb. Cortex* 8, 510–523
- 65 Miyashita, Y. *et al.* (1993) Configurational encoding of complex visual forms by single neurons of monkey temporal cortex. *Neuropsychologia* 31, 1119–1131
- 66 Sakai, K. and Miyashita, Y. (1994) Neuronal tuning to learned complex forms in vision. *Neuroreport* 5, 829–832
- 67 Peissig, J.J. *et al.* (2007) Effects of long-term object familiarity on event-related potentials in the monkey. *Cereb. Cortex* 17, 1323–1334
- 68 Op de Beeck, H.P. *et al.* (2007) Effects of perceptual learning in visual backward masking on the responses of macaque inferior temporal neurons. *Neurosci.* 145, 775–789
- 69 Op de Beeck, H.P. *et al.* (2008) A stable topography of selectivity for unfamiliar shape classes in monkey inferior temporal cortex. *Cereb. Cortex* 18, 1676–1694
- 70 Henson, R. *et al.* (2000) Neuroimaging evidence for dissociable forms of repetition priming. *Science* 287, 1269–1272
- 71 van der Linden, M. *et al.* (2008) Birds of a feather flock together: experience-driven formation of visual object categories in human ventral temporal cortex. *PLoS ONE* 3, e3995
- 72 Erickson, C.A. *et al.* (2000) Clustering of perirhinal neurons with similar properties following visual experience in adult monkeys. *Nat. Neurosci.* 3, 1143–1148
- 73 Harley, E.M. *et al.* (2009) Engagement of Fusiform Cortex and Disengagement of Lateral Occipital Cortex in the Acquisition of Radiological Expertise. *Cereb. Cortex*
- 74 Moore, C.D. *et al.* (2006) Neural mechanisms of expert skills in visual working memory. *J. Neurosci.* 26, 11187–11196
- 75 Rhodes, G. *et al.* (2004) Is the fusiform face area specialized for faces, individuation, or expert individuation? *J. Cog. Neurosci.* 16, 189–203
- 76 Ullman, S. *et al.* (2002) Visual features of intermediate complexity and their use in classification. *Nat. Neurosci.* 5, 682–687
- 77 Vogels, R. and Orban, G. (1994) Does practice in orientation discrimination lead to changes in the response properties of macaque inferior temporal neurons? *Eur. J. Neurosci.* 6, 1680–1690
- 78 Dickey, A.S. *et al.* (2009) Single-unit stability using chronically implanted multielectrode arrays. *J. Neurophysiol.* 102, 1331–1339
- 79 Tolias, A.S. *et al.* (2007) Recording chronically from the same neurons in awake, behaving primates. *J. Neurophysiol.* 98, 3780–3790
- 80 Bondar I.V. *et al.* (2009) Recording chronically from the same neurons in awake, behaving primates. *PLoS ONE*, in press.
- 81 Gibson, J.J. and Radner, M. (1937) Adaptation, aftereffect and contrast in the perception of tilted lines. *J. Exp. Psychol.* 20, 453–467
- 82 Karni, A. and Sagi, D. (1993) The time course of learning a visual skill. *Nature* 365, 250–252
- 83 Schoups, A.A. *et al.* (1995) Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularit. *J. Physiol.* 483 (Pt 3), 797–810
- 84 Karni, A. *et al.* (1994) Dependence on REM sleep of overnight improvement of a perceptual skill. *Science* 265, 679–682
- 85 Censor, N. *et al.* (2006) A link between perceptual learning, adaptation and sleep. *Vision Res.* 46, 4071–4074
- 86 Li, L. *et al.* (1993) The representation of stimulus familiarity in anterior inferior temporal cortex. *J. Neurophysiol.* 69, 1918–1929
- 87 Vuilleumier, P. *et al.* (2002) Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nat. Neurosci.* 5, 491–499
- 88 Sayres, R. and Grill-Spector, K. (2006) Object-selective cortex exhibits performance-independent repetition suppression. *J. Neurophysiol.* 95, 995–1007
- 89 Gauthier, I. and Tarr, M.J. (1997) Becoming a “Greeble” expert: exploring mechanisms for face recognition. *Vision Res.* 37, 1673–1682
- 90 Gold, J. *et al.* (1999) Signal but not noise changes with perceptual learning. *Nature* 402, 176–178
- 91 Goldstone, R.L. *et al.* (2001) Altering object representations through category learning. *Cognition* 78, 27–43
- 92 Grill-Spector, K. *et al.* (2000) The dynamics of object-selective activation correlate with recognition performance in humans. *Nat. Neurosci.* 3, 837–843
- 93 Furmanski, C.S. and Engel, S.A. (2000) Perceptual learning in object recognition: object specificity and size invariance. *Vision Res.* 40, 473–484
- 94 Grill-Spector, K. *et al.* (2006) Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 14–23
- 95 Chiu, C. and Weliky, M. (2004) *The role of neural activity in the development of orientation selectivity*, The MIT Press 117–125
- 96 Mitchell, D.E. (2004) The effects of selected forms of early visual deprivation on perception. In *The Visual Neurosciences* (Chalupa, L.M. and Werner, J.S., eds), pp. 189–204, The MIT Press
- 97 Rodman, H.R. *et al.* (1993) Response properties of neurons in temporal cortical visual areas of infant monkeys. *J. Neurophysiol.* 70, 1115–1136
- 98 Kiorpes, L. and Movshon, J.A. (2004) Neural limitations on visual development in primates. In *The Visual Neurosciences* (Chalupa, L.M. and Werner, J.S., eds), pp. 159–173, The MIT Press
- 99 Sugita, Y. (2009) Innate face processing. *Curr. Opin. Neurobiol.* 19, 39–44
- 100 Pascalis, O. *et al.* (2005) Plasticity of face processing in infancy. *Proc. Natl. Acad. Sci. U. S. A.* 102, 5297–5300
- 101 Sugita, Y. (2008) Face perception in monkeys reared with no exposure to faces. *Proc. Natl. Acad. Sci. U. S. A.* 105, 394–398
- 102 Maurer, D. *et al.* (2005) Missing sights: consequences for visual cognitive development. *Trends Cogn. Sci.* 9, 144–151
- 103 Grill-Spector, K. *et al.* (2008) Developmental neuroimaging of the human ventral visual cortex. *Trends Cogn. Sci.* 12, 152–162